Harvard University
Harvard University Biostatistics Working Paper Series

Year 2008 *Paper* 89

Limitations of Remotely-sensed Aerosol as a Spatial Proxy for Fine Particulate Matter

Christopher J. Paciorek* Yang Liu†

*Harvard School of Public Health, paciorek@hsph.harvard.edu
†Harvard School of Public Health

Limitations of Remotely-sensed Aerosol as a Spatial Proxy for Fine Particulate Matter

Christopher J. Paciorek and Yang Liu

Abstract

Recent research highlights the promise of remotely-sensed aerosol optical depth (AOD) as a proxy for ground-level PM2.5. Particular interest lies in the information on spatial heterogeneity potentially provided by AOD, with important application to estimating and monitoring pollution exposure for public health purposes. Given the temporal and spatio-temporal correlations reported between AOD and PM2.5 , it is tempting to interpret the spatial patterns in AOD as reflecting patterns in PM2.5 . Here we find only limited spatial associations of AOD from three satellite retrievals with PM2.5 over the eastern U.S. at the daily and yearly levels in 2004. We then use statistical modeling to show that the patterns in monthly average AOD poorly reflect patterns in PM2.5 because of systematic, spatially-correlated error in AOD as a proxy for PM2.5 . Furthermore, when we include AOD as a predictor of monthly PM2.5 in a statistical prediction model, AOD provides little additional information to improve predictions of PM2.5 when included in a model that already accounts for land use, emission sources, meteorology and regional variability. These results suggest caution in using spatial variation in AOD to stand in for spatial variation in ground-level PM2.5 in epidemiological analyses and indicate that when PM2.5 monitoring is available, careful statistical modeling outperforms the use of AOD.

Limitations of Remotely-sensed Aerosol as a Spatial Proxy for Fine Particulate Matter

Christopher J. Paciorek, Department of Biostatistics, Harvard School of Public Health
and
Yang Liu, Department of Environmental Health, Harvard School of Public Health

November 3, 2008

Abstract

Recent research highlights the promise of remotely-sensed aerosol optical depth (AOD) as a proxy for ground-level PM2.5. Particular interest lies in the information on spatial heterogeneity potentially provided by AOD, with important application to estimating and monitoring pollution exposure for public health purposes. Given the temporal and spatio-temporal correlations reported between AOD and $PM_{2.5}$, it is tempting to interpret the spatial patterns in AOD as reflecting patterns in $PM_{2.5}$. Here we find only limited spatial associations of AOD from three satellite retrievals with $PM_{2.5}$ over the eastern U.S. at the daily and yearly levels in 2004. We then use statistical modeling to show that the patterns in monthly average AOD poorly reflect patterns in $PM_{2.5}$ because of systematic, spatially-correlated error in AOD as a proxy for $PM_{2.5}$. Furthermore, when we include AOD as a predictor of monthly $PM_{2.5}$ in a statistical prediction model, AOD provides little additional information to improve predictions of $PM_{2.5}$when included in a model that already accounts for land use, emission sources, meteorology and regional variability. These results suggest caution in using spatial variation in AOD to stand in for spatial variation in ground-level $PM_{2.5}$ in epidemiological analyses and indicate that when $PM_{2.5}$ monitoring is available, careful statistical modeling outperforms the use of AOD.

1 Introduction

Epidemiological studies provide evidence that chronic exposure to particulate matter (PM) is related to increased mortality and morbidity (Dockery et al. 1993; Pope et al. 2002; Miller et al. 2007). Studies of the chronic health effects of PM rely on spatial heterogeneity in PM concentrations to identify the effects. A combination of spatial statistical modeling and land use regression can improve estimation of concentrations at fine scales by using land use and meteorological regressors to help estimate concentrations at locations far from monitors (Paciorek et al. 2008b; Yanosky et al. 2008), but efforts still suffer from the sparse spatial representation in the monitoring network.

Remote sensing holds promise for adding spatial information for exposure estimation, particularly in suburban and rural areas far from monitors (e.g., Fig. 1). Satellite-derived aerosol optical depth (AOD) is correlated with ground level PM2.5 (Wang and Christopher 2003; Engel-Cox et al. 2004; Liu et al. 2005;

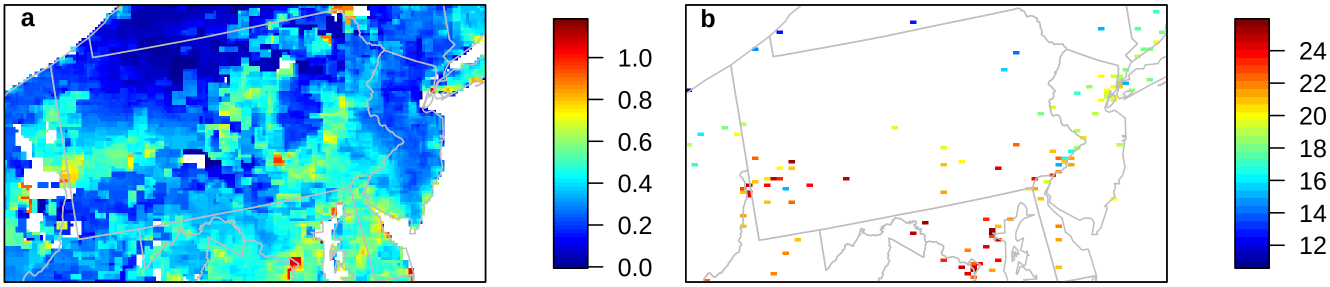

Figure 1. Example of monthly average MODIS AOD and ground-level $PM_{2.5}$: July 2004 in our mid-Atlantic study region of the U.S. Monthly average MODIS AOD (a) and monthly average $PM_{2.5}$ from ground-level monitors (b).

Koelemeijer et al. 2006; Liu et al. 2007; Pelletier et al. 2007; Paciorek et al. 2008a). These correlations occur despite the vertical mismatch between total column aerosol, as measured by AOD, and ground-level $PM_{2.5}$, the level of interest for health studies and the temporal mismatch between 24-hour average PM and daytime (often single snapshot) AOD. These results and success in using AOD to document pollution episodes (e.g., Wang and Christopher 2003; Al-Saadi et al. 2005) have led to excitement about using AOD as a proxy, standing in for PM, or in combination with ground measurements to better predict PM. Our attention focuses on improving empirical prediction, rather than physical explanation, of the spatial patterns of PM.

Most studies of the AOD-PM association focus on temporal (longitudinal) correlations or do not distinguish spatial (cross-sectional) from temporal correlations, but for chronic exposure, estimating spatial heterogeneity is critical. Correlations of long-term averages using matched daily (or hourly) values (e.g., van Donkelaar et al. 2006) do not take into account the large amounts of unavailable retrievals, because of orbit patterns, cloud cover and surface reflectivity, that may seriously compromise the association between available AOD and long-term average PM concentrations. Finally, but critically, simple correlations of AOD and PM do not tell us if AOD improves predictions within a statistical model that already uses information on meteorology, land use, and regional variation, and we are not aware of any such analysis of the use of AOD for exposure assessment.

Here we report both raw empirical results and statistical modeling of the relationship between AOD and PM and the ability of AOD retrievals to improve predictions of ground-level PM in the eastern U.S. Our perspective is a public health one, in which good estimates of PM concentrations are needed over an entire specified spatial region and time period as an input for epidemiological analysis. We first show positive, but moderate and variable, correlations at various temporal scales (Section 3.1). Correlations do not improve when looking at longer-term averages over all the days in a period of time. We introduce a statistical model that treats AOD as proxy data for PM and estimates a PM prediction surface that reflects both the PM and AOD data (Section 3.2). This model shows high sensitivity to assumptions about the structure of the error in AOD as a proxy for PM. The results suggest there are systematic, spatially-correlated differences between AOD and PM and that AOD be disregarded in predicting PM. We confirm this using a simpler model with PM data as the gold standard, regressing PM on AOD and numerous other predictors, showing no gain in predictive power from the use of AOD in an already successful prediction model (Section 3.3).

2 Methods

2.1 Data

All analyses are for the year 2004. Associations of AOD and $PM_{2.5}$ are weak in the western U.S. (Engel-Cox et al. 2004; Liu et al. 2005; Paciorek et al. 2008a), so we focus on the eastern U.S. Our daily exploratory analyses use data east of 100 W longitude. To limit computations with large remote sensing datasets, our longer-term analyses, including the statistical modeling, focus further on a mid-Atlantic region encompassing Pennsylvania, New Jersey and parts of surrounding states (see Fig. 1), which contains the major cities of New York, Philadelphia, Washington DC, Baltimore and Pittsburgh, and their extensive suburban areas, as well as large rural areas in the north. The heterogeneity in population density and the presence of large point source emissions from power plants and industrial plants in the southwestern portion of the region provide a test region with important sources of heterogeneity in pollution.

We use AOD retrievals from three satellite instruments: MODIS (the moderate resolution imaging spectroradiometer), MISR (multi-angle imaging spectroradiometer) and the GOES (Geostationary Operational Environmental Satellite) aerosol/smoke product (GASP). The MODIS and MISR instruments are both aboard the Terra satellite platform, whose polar orbit gives full coverage of the globe at regular intervals, starting in March 2000. Both MISR (at 558 nm) and MODIS (at 550 nm) provide retrievals of AOD, a dimensionless measure of light extinction over the entire vertical column of air through the atmosphere (also known as aerosol optical thickness (AOT)), available through NASA. AOD generally ranges from zero to five, with values greater than one associated with heavy haze. MISR AOD retrievals are at a nominal spatial resolution of 17.6 km with retrievals in the northeast U.S. every 4-7 days depending on location (Liu et al. 2005). MODIS provides AOD retrievals at a nominal resolution of 10 km with each location covered every 1-2 days (Wang and Christopher 2003; Engel-Cox et al. 2004). For a given location in the eastern U.S., the retrievals represent AOD at a single constant time point (10:30-10:45 a.m. local time) because of satellite orbit periodicity. AOD cannot be measured below clouds, so cloud filtering algorithms based on measurements in the infrared portion of the spectrum are employed to detect and omit obscured observations (Engel-Cox et al. 2004). Errors and uncertainties in the cloud filtering can lead to erroneous AOD retrievals, and high surface reflectivity can also prevent retrievals.

GASP AOD (interpolated at 550 nm) is available from GOES-12 (East) imager data, provided by the U.S. National Oceanic and Atmospheric Administration (NOAA) (Prados et al. 2007). GASP AOD is at a nominal spatial resolution of 4 km, but retrievals are less precise than MODIS or MISR AOD because of the coarse spectral resolution and fixed viewing geometry of the sensor. Retrievals are attempted every half-hour during daylight, 10:45-23:45 UTC, but again, cloud cover and high surface reflectivity lead to many missing observations. We use daily average GASP AOD, regardless of the number of retrievals.

We use 24-hour average gravimetric (Federal Reference Method (FRM)) measurements from the US EPA Air Quality System (AQS) with parameter 88101, omitting a small number of IMPROVE monitors, which tend to be placed where few people live. While hourly data are better matched in time to the MODIS and MISR snapshots, the number of hourly monitors is limited and there is no FRM for hourly measurements, plus our interest is in the relationship of PM and AOD at monthly and yearly periods.

In our statistical modeling we use Geographic Information System (GIS)-based and meteorological covariates to help explain PM variation, following Yanosky et al. (2008). Covariates that may help predict PM at fine spatial scale include distance to major roads in three road classes (A1, A2, and A3). We also have point locations of year 2002 primary PM2.5 emissions from EPA's 2002 National Emissions Inventory (NEI). Other covariates are calculated using a GIS at the resolution of the four km grid used in our statistical prediction modeling. These include road density in the three road classes, population density, and elevation at the cell centroid. As a measure of non-point source emissions in each cell, we assign it the density (total emissions divided by county area) of the NEI area-level year 2002 primary PM2.5 emissions in the county of the cell centroid. Meteorological variables are based on the North American Regional Reanalysis (NARR) (Mesinger et al. 2006) fields, available at 32 km resolution every 3 hours. For each 3-hour value and each grid cell, we compute an inverse distance weighted average of the NARR values from the four nearest NARR points to each grid cell centroid. Values are then averaged to the month. Our second statistical model uses wind speed and temperature, but we also considered relative humidity, planetary boundary layer (PBL) height, mean sea-level pressure, and precipitation.

We also make use of a calibrated AOD variable (Paciorek et al. 2008a). The approach accounts for systematic effects of planetary boundary layer (PBL) height, relative humidity, season, and time-invariant regional variation that modify and obscure the relationship between daily PM and daily AOD. The calibration is done by regressing daily AOD values from 2004 across the eastern U.S on daily PM and the variables just mentioned, matched in space and time. By including time-invariant regional variation, we cause the long-term average AOD and long-term average PM to more closely match at large spatial scales, necessarily increasing correlations of PM and AOD. Our hope in including this spatial term is to adjust for large-scale differences between AOD and PM and thereby allow us to exploit common patterns of AOD and PM at smaller spatial scales, to the extent that they exist.

2.2 Exploratory Analyses

Our goal in the exploratory data analysis is to understand the association between AOD and PM at different temporal aggregations to assess the potential of AOD to help predict chronic PM exposure. We start by considering PM and AOD associations at the daily level when AOD and PM are matched such that both are available for a given day and location, mirroring analyses in the published literature. We matched available PM2.5 24-h averages with AOD retrievals from the nearest pixel for each of the three satellite instruments, omitting a small number of monitors for which the nearest pixel centroid is closer to another monitor. Our interest is in fine resolution estimation of $PM_{2.5}$, so unlike other analyses we use individual pixels instead of aggregating AOD across adjoining pixels.

When considering prediction of long-term average PM for chronic epidemiological analyses, the missingness of AOD retrievals means that one relies on a limited subsample of days (determined by weather conditions that also affect PM levels) with AOD retrievals to estimate monthly or yearly pollution. For our mid-Atlantic focus region, MODIS, MISR, and GOES retrievals are available over land on average 16%, 4%, and 38%, respectively, of the days in 2004. Also, for MODIS and MISR, the occurrence of AOD snapshots at the same time every day may not well match daily average pollution. To assess the long-term relationship of AOD and PM, we considered associations of yearly PM and AOD, relating average AOD from available retrievals to average PM based on all available PM monitoring, not just PM data that are matched by day to AOD retrievals. These associations reflect spatial (cross-sectional) associations, thereby eliminating temporal (longitudinal) correlation within a site at the daily level that can obscure the spatial associations. However, simple yearly averaging does not account for the differential frequency of successful AOD retrievals over the seasons in the year (which overweights summer AOD values) nor allow us to consider monthly associations, so we consider working at the monthly level in the Supplementary Material.

2.3 Statistical Modeling

The exploratory analyses do not account for complications such as differing numbers of PM observations and AOD retrievals by location and very fine-scale heterogeneity in PM. Most importantly, correlations of AOD and PM may reflect variability in PM that could be predicted by other sources of information, such as land use or meteorological variables or estimation of large-scale regional variation through spatial smoothing of monitored values, so they may overstate the usefulness of AOD as a predictor in light of other readily available information. To address these issues we turn to formal statistical modeling, analyzing the mid-Atlantic region. Both models are specified in a Bayesian context and are fit by standard Markov chain Monte Carlo methods. We do not use MISR because of the limited number of retrievals.

2.3.1 AOD and PM data as equivalent sources of information

Recent statistical efforts have focused on combining multiple sources of information by treating the sources as reflecting a true, unknown spatial process (Fuentes and Raftery 2005; Gelfand and Sahu 2008; McMillan et al. 2008). Accordingly, we fit statistical models for individual months in which PM and AOD are considered to be separate data sources that reflect the unknown PM surface for a given month. The first stage of the model contains two likelihood terms representing the probabilistic relationship of the PM and AOD data to the underlying processes and covariates. For PM, for an individual month, we specify the likelihood,

$$\mathrm{PM}_i = y_i \quad \sim \quad \mathcal{N}(P_{s(i)} + \sum_k f_k(z_{k,i}), \sigma^2_{y,i}) \tag{1}$$

where the core of the model is the unknown true pollution surface that we want to estimate, which we represent at the scale of a four km grid as P_s, where s indexes grid cells. We represent the monthly average of available 24-h concentration measurements in terms of the gridded pollution surface, locating individual observations, y_i, indexed by location i, within the appropriate grid cell, $s(i)$. $f_k(z_{k,i})$ are smooth regression functions that reflect the effects of local covariates, z_k, that affect PM at scales below four km, a decomposition similar to that of Beelen et al. (2007). In particular, we use distance to the nearest A1 and A2 roads, forcing the effect to be zero beyond 500 m (Zhou and Levy 2007). By modeling the effect of nearby roads (and point emissions - see below), we attempt to account for differences between AOD and PM caused by fine-scale heterogeneity captured by PM monitors but smoothed over in the AOD pixel-level values. Thus we assess the ability of AOD to capture spatial heterogeneity in PM at small scales (10s of kilometers) but not extremely fine scales (meters to kilometers). $\sigma^2_{y,i}$ reflects various components of uncorrelated error, including heteroscedastic components that vary by location based on the number of observations. The likelihood term for AOD is presented below.

The unknown pollution process at the scale of the grid is represented as

$$P_s = \mu + \sum_k h_k(w_{k,s}) + g_s \tag{2}$$

where μ is an overall mean and $h_k(w_{k,s})$ are smoothed regression functions of grid-scale covariates, described below. g_s is a smooth spatial term representing residual spatial structure unaccounted for by covariates, in particular regional variation. The w_k covariates are the density of A1, A2, and A3 roads, population density, elevation and non-point source area emissions. We do not use meteorological covariates here because we fit the model for a single month at a time for simplicity. For individual months the meteorological covariates tend to be spatially smooth, and their influence would be difficult to separate from the smooth spatial term, g_s; omitting them causes their influence to be reflected in the estimate of g_s. Also included the model is a smoothing term that accounts for the effect of point emissions within 100 km, where the effect declines with distance and is estimated from the data within the model fitting. This term is used both as a covariate affecting the individual PM observations based on the point location of each monitor (in (1)) and as a covariate affecting the grid cell scale process, P_s (based on averaging over a subgrid of 16 points within each cell).

We specify the AOD retrievals in an individual month as reflecting the unknown PM process,

$$\text{AOD}_s = a_s \sim \mathcal{N}(\beta_0 + \beta_1 P_s + f_{\text{cloud}}(z_{\text{cloud},s}), \sigma^2_{a,s}), \tag{3}$$

up to additive, β_0, and multiplicative, β_1, bias, with a smooth regression function of cloud cover, $f_{\text{cloud}}(z_{\text{cloud},s})$, where $z_{\text{cloud},s}$ is the monthly average proportion of cloud-free retrievals in the cell, based on the GOES cloud retrieval algorithm. This is included to help account for bias from retrievals systematically missing because of clouds (Koelemeijer et al. 2006; Paciorek et al. 2008a).$\sigma^2_{a,s}$ reflects various components of uncorrelated error, including heteroscedastic components that vary by location based on the number of retrievals. A complicating factor is that for different satellite orbits on different days, the MODIS pixels shift spatially. To account for this, we consider the overlap of all the pixels in an orbit with the four km grid, assigning to each grid cell, s, the value of the MODIS pixel in which the cell centroid falls. Taking the retrievals assigned to each cell, we then average to the monthly level for each cell, giving a_s. For GOES the pixel locations are constant over time, so we average to the monthly level and then assign each grid cell the weighted average AOD of the GOES pixels that the cell overlaps, weighted by the area of overlap. While simplistic, we believe these approaches cause minimal distortion in the AOD values used in the modeling, because of the reasonably smooth local variation in AOD values from pixel to pixel on a daily basis. Note that in this model, we assume any difference between AOD and PM is spatially-uncorrelated noise, which causes estimation of P_s to reflect the spatial structure in both PM and AOD observations.

However, maps of monthly average AOD show strong spatial structure (e.g., Fig. 1a) with limited spatially-uncorrelated noise (i.e., white noise) apparent. This spatial structure may be caused in part by systematic, spatially-correlated differences between AOD and PM, rather than reflecting spatial structure in ground-level PM. Reasons for such structure include spatial structure in pollution aloft above the boundary layer, spatially-patterned biases due to errors in the estimated ground reflectivity, and daily spatial patterns of missing retrievals from clouds with aggregate effect at the monthly level. We will refer to such differences as systematic error. Systematic error could be substantial and it, rather than pixel-scale white noise, may be the dominant factor explaining low correlations with PM seen in our exploratory analyses (Section 3.1). Models that treat AOD as a proxy for PM without accounting for the possibility of systematic error may predict spatial patterns of PM that do not match reality. We assess sensitivity to assumptions about systematic error by including an additive spatial bias term, ϕ_s, represented on the grid scale, to augment the additive bias in (3). Models that include such a term allow for the possibility that AOD retrievals are telling us about systematic (i.e., spatially-correlated) processes specific to AOD that do not reflect spatial patterns in ground-level PM. We estimate ϕ_s using a penalized thin plate spline approach that penalizes complex spatial surfaces, thereby favoring simple surfaces if the data can be sufficiently well-explained by a smooth surface (Ruppert et al. 2003). Such an approach is also used for the other smooth terms in the model, fit naturally within the Bayesian context with the level of smoothing determined by the data. For computational reasons and because the key results are best visualized in model fits of individual months, we fit the model separately for each of the 12 months.

2.3.2 AOD as a Predictor of PM

We also consider a model in which AOD is used as a predictor on the right-hand side of a regression-style model, treating the PM data as the gold standard. This has the benefit of directly calibrating PM to AOD and, if there is little empirical association, discounting AOD as a predictor of PM.

In this model, PM observations are modeled as in (1)

$$\mathrm{PM}_{it} = y_{it} \quad \sim \quad \mathcal{N}(P_{s(i),t} + \sum_k f_k(z_{k,i}), \sigma^2_{it}) \tag{4}$$

while the unknown smooth $\mathrm{PM}_{2.5}$ process, $P_{s,t}$, is similar to (2), but includes AOD, $A_{s,t}$, as a predictor:

$$P_{s,t} \quad = \quad \mu_t + \beta_{1,t} A_{s,t} + \sum_k h_k(w_{k,s,t}) + g_{s,t} \tag{5}$$

This model is fit simultaneously to all 12 months, indexed by the t subscript. For simplicity, we assume that $g_{s,t}$, the residual spatial structure, is not correlated over time. Previous work suggests month to month correlation is limited and that including correlation would do little to improve predictions (Paciorek et al. 2008b). The simplifying assumption makes for easier computations and does not detract from our main point, which is to assess the ability of AOD to improve PM predictions. We allow $\beta_{1,t}$ to vary in an unstructured way with time in case the relationship of AOD and PM varies by season (Paciorek et al. 2008a). In addition to

AOD, and based on some limited variable selection, we include population density, elevation, area emissions, point emissions, and density of A3 roads as GIS-based covariates and wind speed and temperature as meteorological predictors. We also include monthly average cloudiness (as one of the h_k terms) to help account for bias from missing AOD retrievals.

For this approach, a downside is that we require AOD values at all locations. We use the Markov random field approximation to a thin plate spline described in the Supplementary Material to smooth the observed AOD retrievals and make predictions, $A_{s,t}$, at unobserved locations. Pre-processing of pixel-level AOD values to align with the four km grid is as described in Section 2.3.1.

3 Results

3.1 Exploratory analyses

Correlations between daily PM and AOD (matched by day and location) that reflect both temporal and spatial associations are higher than correlations for individual days, taken across spatial locations (Table 1), which reflect only spatial associations. The spatio-temporal associations roughly match those seen in the literature that have been used as evidence of the potential of AOD as a proxy for PM (e.g., Engel-Cox et al. 2004; Liu et al. 2005; Paciorek et al. 2008a). Using AOD directly does not account for meteorological factors and systematic temporal and spatial variability that modify the relationship between AOD and PM, so we also considered the calibrated version of AOD, which somewhat improved the correlations (Table 1).

Table 1 shows near-zero correlations of yearly average PM from all available 24-h values (everyday or every third day sampling) with AOD from available retrievals. Note that for monitors reporting only every three days, missing PM values contribute to noise in the associations seen here. After calibration, AOD is moderately correlated with PM. The calibration includes an overall spatial term, adjusting for any large-scale regional mismatch between AOD and PM that is consistent over the year. This term is responsible for much of the increase in correlation after calibration as it necessarily causes the large-scale patterns of long-term average AOD and PM to more closely match. Our hope is that correcting for such large-scale mismatch allows us to explore whether there is independent information in AOD for predicting smaller-scale patterns of PM, a question that will be answered in the statistical modeling. See the Supplementary Material for results at the monthly level.

3.2 Using AOD as proxy data: Sensitivity to systematic error

For July 2004 for MODIS AOD, Fig. 2 shows model-based predictions of PM and estimates of the spatial bias, ϕ_s, based on (1-3), allowing different amounts of complexity in ϕ_s. When the model omits the spatial bias term (row 1), representing AOD as reflecting PM up to simple additive and multiplicative bias, predictions of PM strongly track AOD spatial patterns (i.e., Fig. 1a). As spatial bias is introduced (row 2) and more flexibility in the spatial bias term is allowed (row 3), predictions increasingly track the PM observations (i.e., Fig. 1b) and results from a model fit without AOD (row 4). The penalized spline model does not stabilize on a model fit with a stable smooth bias surface. When the bias term is forced to be smooth,

Table 1. Correlations of raw and calibrated daily AOD with matched 24-h PM in 2004 for the eastern U.S. (top portion) and correlations of raw and calibrated yearly-average AOD with yearly-average PM (sites with at least 100 daily PM observations, matched in space to AOD) for our mid-Atlantic focal region in 2004 (bottom portion). Yearly averages reflect all available AOD retrievals and all available 24-h average PM concentrations. Calibrated AOD has been adjusted to account for the effects of planetary boundary layer (PBL) height, relative humidity (RH), season, and regional variation in modifying the relationship between daily AOD and PM. Yearly results exclude one site with high PM levels that is outside Pittsburgh and downwind of a major industrial facility.

	Raw AOD			Calibrated AOD		
	MODIS	MISR	GOES	MODIS	MISR	GOES
	Daily values, eastern U.S.					
Temporal plus spatial variation: Overall correlation of daily values across all sites and days.	0.60	0.50	0.38	0.64	0.57	0.40
Spatial variation only: Average of daily spatial correlations (only days with at least 20 matched sites) .	0.35	0.30	0.23	0.45	0.32	0.29
	Yearly averages, mid-Atlantic focal region					
Spatial variation only: Correlation of yearly averages.	0.09	0.25	-0.07	0.49	0.22	0.53

the model is unable to adequately represent the AOD data based on the PM surface and the smooth bias. This suggests there is little common spatial pattern to PM and AOD observations and that true PM is best modeled solely based on ground-level PM data with AOD variability modeled separately. This is seen in row 3 where the model essentially disregards AOD in predicting PM and attributes most of the variability in AOD to ϕ_s. Results for the other 11 months and using GOES AOD or raw AOD give similar conclusions. In summary, systematic error is considerable and critical to include, and predictions are very sensitive to assumptions about this error. If the spatial bias were estimated to be a relatively smooth process, able to be resolved from having PM and AOD data in the same region, the modeling approach provides a means to improve PM prediction by combining the data sources while accounting for the bias. However, these results suggest the bias process is not smooth and cannot be adequately estimated without more dense PM data, which are not available and would largely obviate the need for AOD as a proxy.

3.3 Using AOD as a predictor: Effects on predictive ability

Using AOD as a predictor, predictive ability at both the monthly and yearly resolutions does not improve when either calibrated MODIS or GOES AOD is added to the model already containing the other predictors (Table 2). If we exclude the other predictors (except the GOES cloud term for consistency in comparing the AOD and no-AOD models) and account for spatial variability solely based on spatial smoothing of the observations within the model framework, addition of AOD still shows essentially no improvement in predictions (Table 2). Results are similar when avoiding locations that are most likely affected by very local sources (Table 2). Sensitivity analyses indicate that there was minimal effect of AOD on predictive power when using raw AOD instead, when restricting to monitors in areas with sparse monitoring, or when restricting to everyday monitors (which avoids the extra noise caused by missing monitor values) (not shown). Note that the higher predictability of monthly compared to yearly PM in Table 2 occurs because of the importance of temporal variation, which is easy to estimate based on the monitoring. The lack of improved predictions when using AOD is consistent with the estimates of $\beta_{1,t}$, which are small in magnitude, with wide uncertainty intervals that cover zero, indicating little additional influence of AOD on PM when other variables are in the model.

4 Discussion

We urge caution in assuming that AOD can help improve exposure estimation for PM and particular caution in using AOD to estimate spatial heterogeneity where there is little ground-level PM data for ground-truthing, based on the lack of strong spatial correlation between available AOD retrievals, even after averaging to the monthly or yearly level, and long-term average PM. In a setting in which reasonably dense PM data are available, our statistical modeling results indicate little or no improvement in prediction of long-term average PM when adding AOD. To the extent that raw correlations of AOD and PM reflect the ability of AOD to capture some of the pattern in PM, our results suggest that these can be better estimated by simple spatial smoothing of the available PM data and regression on other predictors, rendering the AOD information extraneous. Koelemeijer et al. (2006) found much stronger correlations of yearly average MODIS AOD

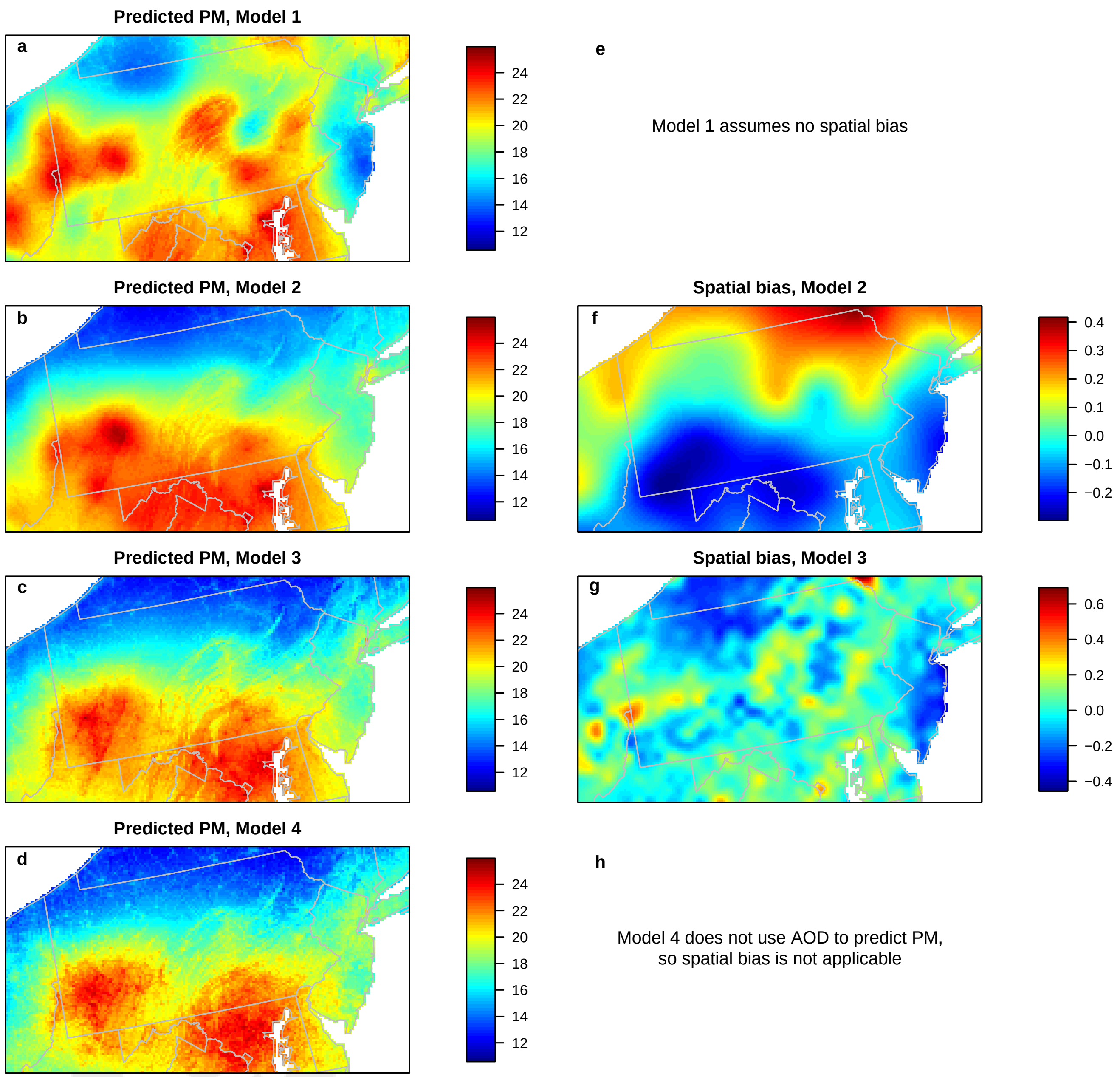

Figure 2. Sensitivity of predicted PM to the characterization of spatial bias. Left column shows PM predictions for models in which AOD and PM observations are treated as data reflecting a common unknown PM process, using calibrated MODIS AOD for July 2004: (a) Model 1: excluding the spatial bias term, ϕ_s, thereby treating AOD as a simple proxy for PM with simple additive and multiplicative bias, (b) Model 2: ϕ_s constrained to be a somewhat smooth process with a maximum of 55 degrees of freedom (a penalized spline with 55 knots), (c) Model 3: ϕ_s relatively unconstrained with a maximum of 755 degrees of freedom, (d) Model 4: AOD not used. Right column shows the corresponding estimated ϕ_s surfaces with the exceptions that for Model 1, ϕ_s is not included in the model (subplot e) and for Model 4, AOD is not used, so ϕ_s is not involved in the model (subplot h).

Table 2. Cross-validation R^2(mean squared prediction error) for yearly average (i.e., space-only variation) and monthly average (i.e., space-time variation) predictions of PM from regression style models with PM data as the gold standard, with and without calibrated AOD and other predictors. Models in the lower half of the table all include GOES cloud cover for consistency in comparing the AOD and no-AOD models. For a given location, only months for which the location has at least four PM daily values are included. Yearly average results include only locations with at least six available months of PM data. The 'population exposure' designation assigned to monitors by EPA indicates that such monitors are not likely to be affected by large, local sources. Results exclude one site with high PM values outside Pittsburgh that is downwind of a major industrial facility.

Model	Yearly averages		Monthly averages	
	All monitors (n=151)	Population exposure monitors (n=130)	All monitors (n=1793)	Population exposure monitors (n=1542)
	Models including land use, emissions, and meteorological predictors			
No AOD	0.580 (1.04)	0.570 (0.93)	0.827 (2.71)	0.839 (2.48)
With calibrated MODIS AOD	0.573 (1.06)	0.564 (0.94)	0.825 (2.73)	0.839 (2.50)
With calibrated GOES AOD	0.572 (1.06)	0.563 (0.95)	0.825 (2.73)	0.838 (2.50)
	Models without land use, emissions, and meteorological predictors			
No AOD	0.463 (1.33)	0.456 (1.18)	0.794 (3.22)	0.810 (2.94)
With calibrated MODIS AOD	0.467 (1.32)	0.459 (1.17)	0.794 (3.22)	0.810 (2.94)
With calibrated GOES AOD	0.467 (1.33)	0.458 (1.17)	0.794 (3.22)	0.810 (2.94)

and PM_{10} in Europe; this may be related to their focus on rural background sites, their larger spatial domain, and the greater variability in their PM concentrations.

Remote sensing is of particular interest in developing countries with little monitoring (e.g., Kumar et al. 2008), but our results on systematic error suggest that spatial patterns seen in AOD may poorly reflect spatial patterns in ground-level PM. Without evidence of strong correlations over space, as opposed to purely temporal correlations, use of AOD to determine spatial heterogeneity in PM may be misleading.

One might ask whether AOD is useful under specific conditions or in specific locations, such as for pollution episodes (e.g., Wu et al. 2006). It is not clear how important such episodes are for long-term average PM prediction or how to include such information only under the circumstances in which it is predictive of PM. To the extent to which AOD is useful in some but not all circumstances, the practical challenge is the need of epidemiologists for exposure estimates without gaps in space or time, often over large domains and long periods of time.

Systematic error such as we see in the satellite proxy for PM can easily be misleading because the spatial structure seen in the proxy leads one to think that the patterns reflect real patterns in the process of interest. In this setting, the evidence suggests that much of this structure does not represent true structure in PM. Such systematic error arises in other contexts (Campbell 1996, p. 378, Robinson 2004, pp. 91-92). It seems likely that deterministic model output used to estimate atmospheric processes, including pollution, such as the widely-used CMAQ model, contain such systematic errors that induce correlated errors in model output, either because of errors in inputs or aspects of the system under study that are not captured by the model.

Some avenues for potential improvement in using remote sensing information to predict PM appear to hold promise. First, Liu et al. (2005) and van Donkelaar et al. (2006) report improvements in relationships of AOD and PM when adjusting for the vertical mismatch based on vertical profile information from an atmospheric chemistry model. However, even after such adjustment, AOD missingness continues to be a problem, and this strategy requires expensive and time-consuming long-term model runs. MISR retrieval information, since it provides reflectance at multiple angles, and ground LIDAR might also prove helpful in distinguishing ground-level aerosol from aerosol aloft. Second, AOD retrieval algorithms aim to accurately estimate AOD, with comparisons to AERONET ground-based observations of AOD. Instead, a tailored approach that modifies AOD retrieval algorithms to directly derive a proxy for ground-level PM may improve upon the current algorithms. The stronger temporal correlations than spatial seen here and in Paciorek et al. (2008a) indicate that when holding surface conditions constant, AOD may be reasonably well-correlated with PM, suggesting that the assumptions about spatial patterns in surface reflectance and particle composition in the retrieval algorithms may be a critical avenue for algorithm improvement, potentially via statistical approaches that are informed by ground-level PM data. Additional work on improving cloud screening algorithms may also be fruitful if it decreases missingness by omitting fewer retrievals not contaminated by clouds or omits retrievals currently suffering from contamination.

Acknowledgements

This work was supported by grant 4746-RFA05-2/06-7 from the Health Effects Institute. The authors thank Shobha Kondragunta and NOAA for access to the GASP AOD retrievals, Steve Melly for GIS processing, Louise Ryan, Charles Stanier, and Helen Suh for comments on the manuscript, and Jeff Yanosky for suggestions. NARR data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, from their Web site at http://www.cdc.noaa.gov.

References

Al-Saadi, J., Szykman, J., Pierce, R., and et al. (2005), Improving national air quality forecasts with satellite aerosol observations *Bulletin of the American Meteorological Society*, 86, 1249–1261.

Beelen, R., Hoek, G., Fischer, P., Brandt, P., and Brunekreef, B. (2007), Estimated long-term outdoor air pollution concentrations in a cohort study *Atmospheric Environment*, 41(7), 1343–1358.

Campbell, J. (1996), *Introduction to Remote Sensing* (2 ed.), New York: The Guilford Press.

Dockery, D., Pope, III, C., Xu, X., Spengler, J., Ware, J., Fay, M., Ferris, B., and Speizer, F. (1993), An association between air pollution and mortality in six U.S. cities *New England Journal of Medicine*, 329, 1753–1759.

Engel-Cox, J., Holloman, C., Coutant, B., and Hoff, R. (2004), Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality *Atmospheric Environment*, 38, 2495–2509.

Fuentes, M. and Raftery, A. (2005), Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models *Biometrics*, 61(1), 36–45.

Gelfand, A. and Sahu, S. (2008), Combining monitoring data and computer model output in assessing environmental exposure Technical report, School of Mathematics, University of Southampton.

Koelemeijer, R., Homan, C., and Matthijsen, J. (2006), Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe *Atmospheric Environment*, 40, 5304–5315.

Kumar, N., Chu, A., and Foster, A. (2008), Remote sensing of ambient particles in Delhi and its environs: estimation and validation *International Journal of Remote Sensing*, 29, 3383–3405.

Liu, Y., Franklin, M., Kahn, R., and Koutrakis, P. (2007), Using aerosol optical thickness to predict ground-level PM2.5 concentrations in the St. Louis area: a comparison between MISR and MODIS *Remote Sensing Environment*, 107, 33–44.

Liu, Y., Sarnat, J., Kilaru, V., Jacob, D., and Koutrakis, P. (2005), Estimating ground-level PM2.5 in the eastern United States using satellite remote sensing *Environmental Science and Technology*, 39, 3269–3278.

McMillan, N., Holland, D., Morara, M., and Feng, J. (2008), Combining numerical model output and particulate data using Bayesian space-time modeling *submitted*.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P., Ebisuzaki, W., et al. (2006), North American Regional Reanalysis *Bulletin of the American Meteorological Society*, 87(3), 343–360.

Miller, K., Siscovick, D., Sheppard, L., Shepherd, K., Sullivan, J., Anderson, G., and Kaufman, J. (2007), Long-term exposure to air pollution and incidence of cardiovascular events in women *New England Journal of Medicine*, 356, 447–459.

Paciorek, C., Liu, Y., Moreno-Macias, H., and Kondragunta, S. (2008), Spatio-temporal associations between GOES aerosol optical depth retrievals and ground-level PM2.5 *Environmental Science and Technology*, 42, 5800–5806.

Paciorek, C., Yanosky, J., Puett, R., Laden, F., and Suh, H. (2008), Practical large-scale spatio-temporal modeling of particulate matter concentrations *Annals of Applied Statistics*, in press.

Pelletier, B., Santer, R., and Vidot, J. (2007), Retrieving of particulate matter from optical measurements: a semiparametric approach *Journal of Geophysical Research*, 112, D06208.

Pope, III, C., Burnett, R., Thun, M., Calle, E., Krewski, D., Ito, K., and Thurston, G. (2002), Lung cancer, cardiopulmonary mortality and long-term exposure to fine particulate air pollution *Journal of the American Medical Association*, 287, 1132–1141.

Prados, A., Kondragunta, S., Ciren, P., and Knapp, K. (2007), GOES aerosol/smoke product (GASP) over North America: comparisons to AERONET and MODIS observations *Journal of Geophysical Research-Atmospheres*, 112, D15201.

Robinson, I. (2004), *Measuring the Oceans from Space: the Principles and Methods of Satellite Oceanography*, Berlin: Springer.

Ruppert, D., Wand, M., and Carroll, R. (2003), *Semiparametric Regression*, Cambridge, U.K.: Cambridge University Press.

van Donkelaar, A., Martin, R., and Park, R. (2006), Estimating ground-level PM2.5 using aerosol optical depth determined from satellite remote sensing *Journal of Geophysical Research*, 111(D21), 21201.

Wang, J. and Christopher, S. A. (2003), Intercomparison between satellite-derived aerosol optical thickness and PM2.5 mass: Implications for air quality studies *Geophysical Research Letters*, 30, 2095.

Wu, J., Winer, A., and Delfino, R. (2006), Exposure assessment of particulate matter air pollution before, during, and after the 2003 Southern California wildfires *Atmospheric Environment*, 40, 3333–3348.

Yanosky, J., Paciorek, C., Schwartz, J., Laden, F., Puett, R., and Suh, H. (2008), Spatio-temporal modeling of chronic PM10 exposure for the Nurses' Health Study *Atmospheric Environment*, 42, 4047–4062.

Zhou, Y. and Levy, J. (2007), Factors influencing the spatial extent of mobile source air pollution impacts: a meta-analysis *BMC Public Health*, 7(1), 89.

Supplementary Material: Limitations of Remotely-sensed Aerosol as a Spatial Proxy for Fine Particulate Matter

Christopher J. Paciorek,* Department of Biostatistics, Harvard School of Public Health
and
Yang Liu, Department of Environmental Health, Harvard School of Public Health

October 31, 2008

S1 Exploratory analysis results based on monthly values

Here we consider spatial associations of monthly average PM and AOD, matched in space, for each of the 12 months of 2004, to supplement the daily and yearly analyses in the main paper. We focus on MODIS and GOES because of the extremely low retrieval availability of MISR, which would require interpolation of AOD values over very large areas with no retrievals.

Even after averaging to the month, some locations have no retrievals, so we use a statistical smoothing model to estimate an AOD surface for each month, using a computationally efficient Markov random field representation of a thin-plate spline (Rue and Held 2005, Sec. 3.4.2; Yue and Speckman, unpublished manuscript) that readily fits the AOD retrieval data for each month (there are as many as 15,000 observations in a month). We use a heteroscedastic residual variance that accounts for the differing number of retrievals in different locations. While the model has the flexibility to do either substantial or little smoothing, in practice because the AOD values are fairly smooth at a fine spatial scale, the smoothed fields look similar to the raw fields, but with imputed values where no data are available.

Correlations are higher in the warmer months, but are moderate at best (Fig. S1). Results are similar when restricting to locations with at least three days with AOD retrievals in a given month. The poor correlations result in part because of the limited retrieval availability, particularly during winter, seen in red in Fig. S1.

Taking the twelve smoothed monthly values and averaging to the year at each location, Table S1 shows near-zero correlations of raw AOD and moderate correlations for calibrated AOD with yearly average PM. By averaging with equal weights for each month, we attempt to account for the differing retrieval availability in different seasons. Focusing on the warm season, to avoid months with few retrievals, correlations of calibrated AOD increase but those of raw AOD do not. If we consider only monitors with at least 300 observations (i.e., monitors reporting daily with little missing data), correlations for the calibrated AOD are

*email: paciorek@hsph.harvard.edu

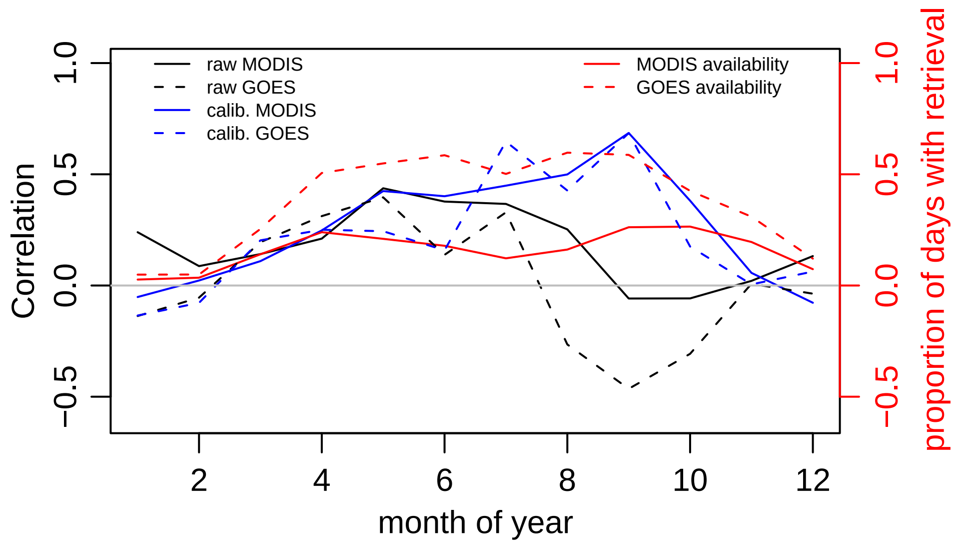

Figure S1: Correlations (across space) of monthly average smoothed AOD with spatially-matched PM by month for the mid-Atlantic focal region for locations with at least 5 daily PM values in the month (black for raw AOD and blue for calibrated AOD). Red lines show the proportion of days with a successful retrieval by month, averaged over non-water areas in the region. For consistency, results exclude the site excluded in Table 1 in the main paper, but this exclusion has little effect here.

similar, while for raw AOD, they are higher but still moderate in magnitude. The increased correlations may be related to the fact that daily monitors are more likely to be in locations with high PM concentrations; daily monitors are more likely to be categorized by EPA as monitors sited to monitor high concentration areas (18% of daily monitors but only 5% of non-daily monitors).

Table S1: Correlations (across space) of yearly and warm-season (April-October) average AOD and spatially-matched PM (sites with at least 100 daily PM observations) for the mid-Atlantic focal region. Averages are computed by averaging over monthly values; for AOD these are produced by spatial smoothing of the available monthly data, thereby filling in missing values at the monthly level. Results exclude the site excluded in Table 1 in the main paper.

	Raw AOD		Calibrated AOD	
	MODIS	GOES	MODIS	GOES
Overall correlation	0.19	-0.06	0.53	0.44
Correlation for April-October	0.05	-0.19	0.65	0.70
Overall correlation, daily monitors	0.43	0.27	0.45	0.40